基于模板的最小能量变形皮层表面重建

Patrick Madlindl¹, Fabian Bongratz^{1,2}, and Christian Wachinger^{1,2}

¹ Lab for AI in Medical Imaging, Technical University of Munich, Munich, Germany {patrick.madlindl,fabi.bongratz}@tum.de

摘要 皮层表面重建(CSR)是从磁共振成像(MRI)中进行神经影像分析的基础,使大脑皮层的形态学研究和功能脑图谱成为可能。基于学习的 CSR 方面的最新进展极大地加速了处理速度,使得通过解剖模板的变形在 几秒钟内完成重建成为可能。然而,确保所学到的变形在变形能量方面是最优的,并且在整个训练过程中保持一致性仍然是一个特别挑战。在这项工作中,我们设计了一个最小能量变形(MED)损失函数,作为对 CSR 中广泛使用的 Chamfer 距离的补充和对变形轨迹的正则化器。我们将它整合到最近的 V2C-Flow 模型中,并展示了在以前被忽视的训练一致性和可重复性方面的显著改进,而不会影响重建精度和拓扑正确性。

Keywords: 皮质表面重建 · 正则化 · 可重复性

1 介绍

皮质表面重建(CSR)是结构和功能性神经影像分析的基石 [5,6]。通过 从磁共振成像(MRI)中提取脑灰质的边界,并将其表示为高分辨率三角面 网格,CSR 能够对高度折叠的皮层片进行拓扑忠实建模。这进一步允许精确的形态测量,这对于调查正常大脑发育 [1] 以及理解神经和精神疾病的发生和发展 [13,16] 至关重要。CSR 在已建立的神经成像工具箱中实现,例如 FreeSurfer [6] 和 CAT12 [8]。然而,这些标准方法的一个关键限制是长时间运行,通常以小时为单位,这使得处理大型数据库变得复杂。

最近,基于深度学习的 CSR 作为一种强大的替代方案出现了,提供了显著的运行时间减少——从几小时缩短到几秒。一种常见的方法是变形一个通用输入脑模板,例如人群平均模板如 FsAverage [7],直接根据从 MRI 数据提取的深度特征转换到个体轮廓中 [3,15,20,19,12]。与基于分割的 [11,17,14] 和隐式 CSR [9,4] 相比,模板化方法避开了昂贵的等值面抽取和拓扑修正,因为皮层条带的球形拓扑结构通过变形一个拓扑正确的模板网格而得以保

² Munich Center for Machine Learning, Munich, Germany

留(直到潜在的自交)。此外,由于与这些分析所依赖的人群模板之间建立了对应关系,下游分析如分区和组比较得到了极大的简化 [3]。在训练用于 CSR 的神经网络时,通常会采用 FreeSurfer [6] 网格作为银标准参考,因为 获取大量合理的 3D 手动描绘皮层边界的难度较大。作为一种重建损失,由于其计算效率和通用性,点云距离是一个受欢迎的选择 [3,15,20,12]。重要的是,点云损失在训练表面上不需要逐点对应。此外,正则化损失函数,如法线一致性或边缘损失,确保输出表面的规则性和光滑性 [3,15]。

然而,尽管基于变形的 CSR 近年来非常流行,先前的研究中却忽视了两个重要的属性:训练结果的可重复性和学习到的变形的最优性。与传统的 CSR 相比,训练过的 CSR 模型由于不同的(通常是未报告的)初始化和训练随机性,在不同次训练之间可能会有所不同,即使数据集和超参数(通常会报告)保持不变。与重复测试评估 [18] 相比,其中相同的模型被应用于同一主体在大致相同时间获取的多次扫描,可重复性关注的是基于不同的模型初始化对训练结果可再现性的关注。确保可重复性对于 CSR 模型及其下游统计分析的可信度至关重要。另一方面,最优性指的是模型学习到的变形轨迹的有效性。如果没有适当的约束条件,这些轨迹可能会变得不必要的复杂或在解剖学上不合理,从而导致次优重建,即使它们仍可能最小化点对点损失如 Chamfer 距离。虽然存在一种基于切片 Wasserstein 距离 [14] 的最优传输方法用于 CSR,但它需要白质分割作为起点,并因此不适用于从人群模板开始的大型变形。

在这项工作中,我们提出了一种新的正则化损失函数,最小能量变形 (MED) 损失,用于基于模板的 CSR。这个损失函数建立在这样的直觉上:从模板到重建的最佳变形应该包含扭曲顶点路径长度上的最小变形能量。概念上讲,MED 损失补充了重建损失,后者衡量预测与参考表面之间的偏差,因为它针对的是从模板到变形预测的变形,如 Figure 1 所示。在我们的实验中,我们展示了这种方法的有效性,它提高了最近提出的 V2C-Flow模型 [3] 预测的可重复性和最优性。

2 方法

我们在 Figure 1 中提供了我们方法的示意概述,该方法被称为最小能量 形变体 Vox2Cortex (V2C-MED)。

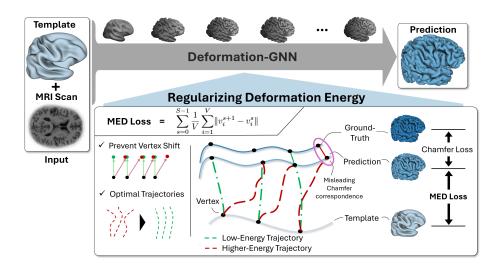


图 1: V2C-MED 和所提出的最小能量变形(MED)正则化损失的示意图概述。与标准的 Chamfer 重建损失相比,该损失衡量预测和地面真实之间的表面差异,MED 损失在预测和输入模板之间引入了一种新颖的互补约束。具体而言,MED 损失计算 V 个顶点间的平均路径长度,促进具有最小变形能量的变形。基于 V2C-Flow[3] 的变形-GNN 预测了 S 步数值积分步骤,从而计算出逐顶点的路径长度。

2.1 V2C-流

V2C-MED 基于 Vox2Cortex-Flow (V2C-Flow) [3],预测一个动态的顶点变形场,将输入的网格模板变形为输入 MRI 扫描中可见的 3D 脑轮廓。该变形场采用前向欧拉积分方案进行集成,总共包含 S=10 个积分步骤。简而言之,V2C-Flow 利用图神经网络 (GNN) 处理表面网格,并使用 3D 卷积神经网络 (CNN) 从 MRI 扫描中提取深度特征。它基于曲率加权的 Chamfer 损失 \mathcal{L}_{Ch} 、网格边缘损失 \mathcal{L}_{e} 、法线一致性损失 \mathcal{L}_{nc} 和交叉熵分割损失 \mathcal{L}_{ce} 的总和进行训练。具体来说,原始的 V2C-Flow 损失函数表示为

$$\mathcal{L}_{\text{V2C-Flow}} = \mathcal{L}_{Ch} + \mathcal{L}_e + 0.001 \,\mathcal{L}_{nc} + \mathcal{L}_{ce}, \tag{1}$$

其中,损失项的相对权重遵循 V2C-Flow 中提出的原始配置。此外, V2C-Flow 通过虚拟边连接内部(白质)和外部(皮层)表面,并同时变形它们以获得一致的输出。关于架构和训练细节,请参考原始的 V2C-Flow 论文 [3]。

2.2 最小能量变形 (MED) 损失

我们最小能量变形(MED)损失的动机来源于这样一个事实,即 V2C-Flow 与其他相关的 CSR 方法 [19,15,20] 一样,仅基于生成的预测进行训练,使用 Chamfer 重建损失,而不考虑如何这些预测是如何获得的,即实际变形轨迹。具体来说,在 Chamfer 损失中,由于次优网络初始化和预测与真实表面之间的误导最近点距离而导致的顶点位移很容易发生。此外,带有边缘和平滑一致性的光滑正则化在此方面也无济于事。同时,从最小变形能量的角度来看,最优变形对于稳定、可重复且可靠的模型预测至关重要。

作为一种解决方案, 我们建议基于所有 (V) 模板顶点的平均 L2 路径长度来正则化训练损失。路径长度是从 S 数值积分步骤的总和中获得, 并在白质 (W) 和脑皮层 (P) 边界之间累积, 得到以下训练损失函数:

$$\mathcal{L}_{MED} = \sum_{c \in \{W,P\}} \sum_{s=0}^{S-1} \frac{1}{V} \sum_{i=1}^{V} ||v_{c,i}^{s+1} - v_{c,i}^{s}||_{2}.$$
 (2)

与基于最优传输/沃瑟斯坦距离的损失函数 [14] 不同,后者需要对预测形状和真实形状之间的最优传输计划进行昂贵的搜索,我们利用模板和变形形状在 Equation (2) 中点到点顶点对应的关系。通过平等对待每个顶点,即给每个顶点一个单位权重, \mathcal{L}_{MED} 提供了全局形变能量的一种度量。在训练过程中,我们将它添加到原始的 V2C-Flow 损失中,得到

$$\mathcal{L}_{\text{V2C-MED}} = \mathcal{L}_{\text{V2C-Flow}} + \lambda \, \mathcal{L}_{MED} \tag{3}$$

作为我们 V2C-MED 模型的损失函数。正则化权重 λ 的调整基于我们的 CSR 验证集,表明设置 $\lambda=0.01$ 是实际中的合理选择。

2.3 可重复性评估

之前,CSR 模型的评估是基于表面准确性、拓扑正确性和测试-重新测试可靠性进行的,并使用了既定指标 [3,9,15,17]。然而,不同模型初始化下的训练结果可重复性问题一直被广泛忽视。为此,我们用随机种子 τ_0,τ_1,τ_2 训练了三个模型,影响了 GNN 和 CNN 模型权重的初始化。在 V2C-Flow 和 V2C-MED 中,除了输出层外,我们将 GNN 权重使用标准正态分布进行初始化,为了训练稳定性,我们始终将输出层初始化为零 [3]。对于 CNN,我们采用了 PyTorch(v1.10)中实现的标准 Kaiming 均匀初始化 [10]。

作为可重复性的一项度量,我们计算每个顶点 v 与其在三次运行中的平均值 \bar{v} 的均方根偏差(RMSD)。设 $v(\tau_i)$ 是用种子 τ_i 初始化的模型中预测的一个顶点,那么我们的逐顶点 RMSD 给定为

$$RMSD(v) = \sqrt{\frac{1}{3} \sum_{i=0}^{2} ||v(\tau_i) - \bar{v}||_2^2}.$$
 (4)

最后,通过聚合测试集中所有顶点和受试者的均值来获得单一的可重复性分数。

2.4 实现与实验设置

我们实现的 V2C-MED 基于官方公开的 V2C-Flow 仓库,该仓库在 https: //github.com/ai-med/Vox2Cortex 上可以找到。为了确保公平且受控的比 较,我们采用了与[3]中描述相同的训练协议,包括数据预处理、网络架构 和优化设置。这种受控设置使我们能够隔离并评估所提出的最小能量形变 (MED) 正则化损失的具体贡献。所有实验均在相同条件下进行,以突出 MED 损失对训练稳定性、可重复性和重建质量的影响。由于其在下游应用 中的受欢迎程度,我们使用了FsAverage模板[7],分别采用平滑版本(称为 V2C-Flow-S 的 [3]) 作为所有模型的输入。我们考虑了两种不同的训练模板 分辨率: ico-6 (每个表面有 40,962 个顶点) 和 ico-7 (每个表面有 163,842 个 顶点); 在测试时, 我们始终使用可能的最高模板分辨率 (ico-7)。先前已经 表明, 在无法获取 VRAM 超过 24GB 的 GPU 的情况下, 测试时的表面放大 是有益的[2]。此外,为了可行性考虑(我们需要针对每个配置从头开始训练 三个模型,并用不同的随机种子来评估可重复性,请参见 Section 2.3),我 们在这里专注于右脑半球。通过在矢状面将原始扫描分割成两半,我们获得 了维度为 [96×208×192] 的输入扫描。我们在每个模型上使用了最多 85 个 训练周期,这在该设置下已经足够收敛。这相当于在单个 Nvidia A100 GPU 上使用 ico-6 分割进行大约两天的训练;对于 ico-7 模板,每个模型的训练 大约需要四天。我们始终选择基于验证分数具有最低重建误差的模型来进行 最终测试集评估。

2.5 数据集

We used data from the Alzheimer's Neuroimaging Initiative (ADNI) (http://adni.loni.usc.edu/), which comprises T1-w MRI scans (1mm isotropic

表 1: 我们报告了基于测试集的重建准确性,包括平均对称表面距离(ASSD)、以自相交面百分比(%SIF)表示的拓扑正确性、测试-重新测试可靠性(TRT)、归一化图像空间中的变形能量以及如 Section 2.3 中定义的均方根偏差(RMSD)的再现性。所有值均为三次训练运行的平均值±标准差,除 RMSD外(在网格顶点上的平均值±标准差)。对于所有指标,数值越低越好;最佳值已突出显示。

		训练 ico-6		训练 ico-7	
Metric	Surf.	V2C-Flow	V2C-MED	V2C-Flow	V2C-MED
准确度 (ASSD, 毫米)	WM Pial	0.200_0.000	0.292 ± 0.010 0.297 ± 0.002		0,0.
%SIF	WM	1.326±0 .116	0.741±0 .033	1.766±0 .733	2.250±0 .303
可靠性(TRT,毫米)	Pial WM		1.476 ± 0.211 1.211 ± 0.028		
	Pial	1.404±0 .039	1.323 ± 0.032	1.312±0 .030	1.202±0 .024
Deformation Energy	WM Pial		0.084 ± 0.001 0.094 ± 0.001		
Reproducibility (RMSD, mm)	WM Pial		1.348 ± 0.388 1.436 ± 0.413		

resolution, registered to MNI152 standard space) of subjects diagnosed with Alzheimer's Disease, mild cognitive impairment, and cognitively normal condition. We only used baseline scans, split into 1154 training, 169 validation, and 323 test cases. We balanced the splits with respect to diagnosis, sex, and age. Moreover, we use a test-retest (TRT) database [18], comprising 40 scans from 3 subjects, respectively, to evaluate the reliability of reconstructions across scans. As a silver-standard reference for the reconstruction accuracy, we used FreeSurfer (v7.2) [6]. Note, however, that the evaluation of the training reproducibility is independent of FreeSurfer.

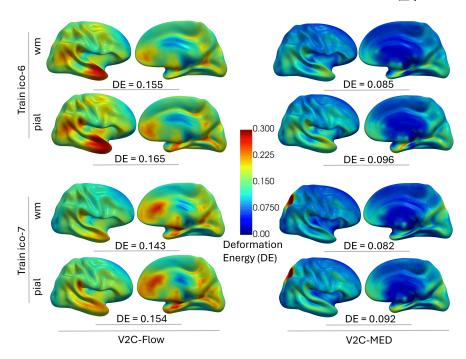


图 2: 在我们的测试集上,V2C-Flow(左)和 V2C-MED(右)训练运行中 ico-6(顶部)和 ico-7(底部)的白质(wm)和皮层表面逐顶点平均变形能量。较低的值表示能量更低的变形。可视化基于平滑后的 FsAverage 模板。

3 结果与讨论

我们报告了所有训练模型在测试集上的表现,如 Table 1 所总结。报告的标准差(SD)反映了每个模型的三次独立训练运行之间的变异性。RMSD是一个例外,它按照设计跨所有三次运行汇总结果;在这种情况下,在网格顶点上计算 SD。

首先,作为一个验证检查,我们观察到所提出的MED 损失是有效的,因为它在所有情况下将变形能量减少了大约50-60% (参见 Figure 2)。同时,重建精度保持特别稳定。尽管最好的精度是在原始的V2C-Flow 模型上获得的,该模型使用 ico-7 网格进行训练,但与 V2C-MED 相比优势很小,V2C-Flow和 V2C-MED 之间的平均精度差异小于0.015毫米。另一方面,训练模板的分辨率对模型准确性的影响要大得多(~为0.2毫米,而使用 ico-6 进行训练时为~0.3毫米),正如预期的那样。就表面精度的可重复性而言,所有模型表现相对较好,这可以通过多次运行中平均对称表面距离(ASSD)的标准

差较低来证明。值得注意的是,在大多数情况下 V2C-MED 显示出了轻微的 优势,除了在使用 ico-6 分辨率训练白质(WM)表面时。然而,RMSD 揭示了 V2C-Flow 和 V2C-MED 之间的显著差异: 当使用 ico-6 分辨率进行训练时,MED 损失将 RMSD 减少了大约 25%,而对于 ico-7 训练则高达约 50%。这表明尽管 V2C-Flow 在多次运行中保持了一致的重建精度,但潜在的变形模式变化很大——强调了如何在没有 MED 正则化的情况下模型达到一定精度的稳定性较差。最后,我们发现当使用 ico-6 模板进行训练时,MED 损失对自相交面的比例和测试重测可靠性有正面影响。总之,这些结果展示了 V2C-MED 在可重复性和变形能量方面的一致性定量改进,同时保持了高度的重建保真度。此外,在不同网格分辨率下的一致改善表明 MED 损失具有良好的泛化能力,并且可以有效地整合到各种训练设置中。

为了进一步了解 CSR 的局部可重复性,我们绘制了每个顶点在 Figure 3 中的 RMSD。这些可视化结果补充了来自 Table 1 的研究发现;它们表明通过 MED 损失函数实现的可重复性改进在整个皮质片上是一致的。然而,在侧顶叶皮层和舌回中,顶点位置仍然存在相当大的差异。

4 结论

我们引入了最小能量变形(MED)损失,并将其集成到 V2C-Flow 中用于皮质表面重建。所得到的皮质表面重建方法,命名为 V2C-MED,在将输入模板扭曲到预测表面所需的变形能量方面表现出显著改进。与之前的研究所不同,我们进行了多次训练,使用不同的随机种子来评估跨不同初始化模型的训练可重复性和预测一致性。我们的结果显示,MED 损失促进了在不同训练场景下的更高可重复性和一致性。这些发现突显了作为增强基于模板的皮质表面重建模型稳定性和可信度的强大正则化策略的 MED 损失潜力。

Acknowledgments. 本研究部分得到了德国研究基金会(DFG,编号 460880779)的支持。我们感谢莱布尼茨超级计算中心(www.lrz.de)提供的计算资源。

参考文献

- 1. Bethlehem, R.A.I., Seidlitz, J., White, S.R., et al.: Brain charts for the human lifespan. Nature **604**(7906), 525 533 (Apr 2022)
- 2. Bongratz, F., Rickmann, A.M., Pölsterl, S., Wachinger, C.: Vox2cortex: Fast explicit reconstruction of cortical surfaces from 3d mri scans with geometric deep

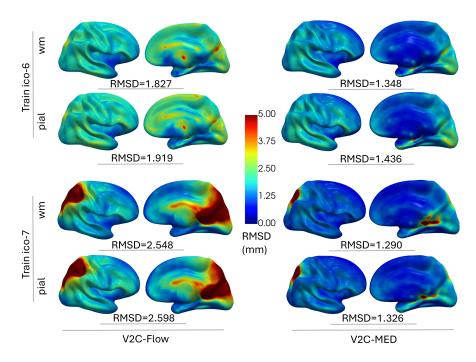


图 3: 顶点-wise 可重复性 (RMSD) 在我们的测试集上的 wm 和 pial 表面,在 ico-6 (顶部) 和 ico-7 (底部) 训练中,针对 V2C-Flow (左侧)和 V2C-MED (右侧)。数值越低表示可重复性越好。可视化基于平滑的 FsAverage 模板。

neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 20773-20783 (Jun 2022)

- Bongratz, F., Rickmann, A.M., Wachinger, C.: Neural deformation fields for template-based reconstruction of cortical surfaces from mri. Medical Image Analysis 93, 103093 (Apr 2024)
- Cruz, R.S., Lebrat, L., Bourgeat, P., Fookes, C., Fripp, J., Salvado, O.: Deepcsr: A 3d deep learning approach for cortical surface reconstruction. In: 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 806–815 (2021). https://doi.org/10.1109/WACV48630.2021.00085
- Esteban, O., Markiewicz, C.J., Blair, R.W., Moodie, C.A., Isik, A.I., Erramuzpe, A., Kent, J.D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S.S., Wright, J., Durnez, J., Poldrack, R.A., Gorgolewski, K.J.: fmriprep: a robust preprocessing pipeline for functional mri. Nature Methods 16(1), 111 116 (Dec 2018), http://dx.doi.org/10.1038/s41592-018-0235-4

- Fischl, B.: Freesurfer. NeuroImage 62(2), 774–781 (2012). https://doi.org/https://doi.org/10.1016/j.neuroimage.2012.01.021, 20 YEARS OF fMRI
- 7. Fischl, B., et al.: High-resolution intersubject averaging and a coordinate system for the cortical surface. Human Brain Mapping 8(4), 272–284 (1999)
- 8. Gaser, C., Dahnke, R., Thompson, P.M., Kurth, F., Luders, E., theăAlzheimer'săDiseaseăNeuroimagingăInitiative: Cat: a computational anatomy toolbox for the analysis of structural mri data. GigaScience 13, giae049 (08 2024). https://doi.org/10.1093/gigascience/giae049
- Gopinath, K., Desrosiers, C., Lombaert, H.: Learning joint surface reconstruction and segmentation, from brain images to cortical surface parcellation. Medical Image Analysis 90, 102974 (Dec 2023). https://doi.org/10.1016/j.media.2023.102974
- He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). p. 1026 1034. ICCV '15, IEEE Computer Society, USA (2015). https://doi.org/10.1109/ICCV.2015.123
- Henschel, L., Conjeti, S., Estrada, S., Diers, K., Fischl, B., Reuter, M.: Fastsurfer

 a fast and accurate deep learning based neuroimaging pipeline. NeuroImage 219,
 117012 (Oct 2020). https://doi.org/10.1016/j.neuroimage.2020.117012
- 12. Hoopes, A., Iglesias, J.E., Fischl, B., Greve, D., Dalca, A.V.: Topofit: Rapid reconstruction of topologically-correct cortical surfaces. In: Konukoglu, E., Menze, B., Venkataraman, A., Baumgartner, C., Dou, Q., Albarqouni, S. (eds.) Proceedings of The 5th International Conference on Medical Imaging with Deep Learning. Proceedings of Machine Learning Research, vol. 172, pp. 508–520. PMLR (Jul 2022), https://proceedings.mlr.press/v172/hoopes22a.html
- Kuperberg, G.R., Broome, M.R., McGuire, P.K., David, A.S., Eddy, M.I., Ozawa,
 F., Goff, D.C., West, W.C., Williams, S.C.R., van der Kouwe, A., Salat, D.H.,
 Dale, A.M., Fischl, B.R.: Regionally localized thinning of the cerebral cortex in schizophrenia. Schizophrenia Research 60, 199–200 (2003)
- 14. Le, T.T., Nguyen, K., shanlin sun, Han, K., Ho, N., Xie, X.: Diffeomorphic mesh deformation via efficient optimal transport for cortical surface reconstruction. In: The Twelfth International Conference on Learning Representations (2024), https://openreview.net/forum?id=gxhR8vUQb
- Lebrat, L., Santa Cruz, R., de Gournay, F., Fu, D., Bourgeat, P., Fripp, J., Fookes, C., Salvado, O.: Corticalflow: A diffeomorphic mesh deformation module for cortical surface reconstruction. In: Advances in Neural Information Processing Systems. vol. 34 (Dec 2021)

- Lerch, A.P., Pruessner, J.C., Zijdenbos, A.P., Hampel, H., Teipel, S.J., Evans, A.C.: Focal decline of cortical thickness in alzheimer's disease identified by computational neuroanatomy. Cerebral Cortex 15(7), 995–1001 (2005). https://doi. org/10.1093/cercor/bhi158
- Ma, Q., Li, L., Robinson, E.C., Kainz, B., Rueckert, D., Alansary, A.: Cortex-ode: Learning cortical surface reconstruction by neural odes. IEEE Transactions on Medical Imaging 42(2), 430–443 (2023). https://doi.org/10.1109/TMI.2022.3206221
- 18. Maclaren, J., Han, Z., Vos, S.B., Fischbein, N., Bammer, R.: Reliability of brain volume measurements: A test-retest dataset. Scientific Data 1(1) (Oct 2014). https://doi.org/10.1038/sdata.2014.37
- Rickmann, A.M., Bongratz, F., Wachinger, C.: Vertex correspondence in cortical surface reconstruction. In: Greenspan, H., Madabhushi, A., Mousavi, P., Salcudean, S., Duncan, J., Syeda-Mahmood, T., Taylor, R. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. pp. 318–327. Springer Nature Switzerland, Cham (2023)
- Santa Cruz, R., Lebrat, L., Fu, D., Bourgeat, P., Fripp, J., Fookes, C., Salvado, O.: CorticalFlow++: Boosting Cortical Surface Reconstruction Accuracy, Regularity, and Interoperability, p. 496 – 505. Springer Nature Switzerland (2022). https://doi.org/10.1007/978-3-031-16443-9_48